Lec. 8 – Expected value/mean/average, variance, standard deviation, median

① Discrete variables

Ex. 1a Consider a simple game: You roll a fair die (a cube) with numbers 1 through 6 on its faces. If it lands on a face with number n, you win $\$n$. If you roll the die 100 times, what amount do you expect to win?

Sol'n: 1)

$\$n$	probability to get $\$n$
1	$p_1 = 1/6$
2	$p_2 = 1/6$
⋮	⋮
6	$p_6 = 1/6$

So the amount you should expect to win after rolling the die once is:

$$\underbrace{\$1\cdot p_1}_{\text{if lands on 1}} + \underbrace{\$2\cdot p_2}_{\text{if lands on 2}} + \dots + \underbrace{\$6\cdot p_6}_{\text{if lands on 6}} = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + \dots + 6\cdot\frac{1}{6} = \$3.5$$

New notation: $E_1(R) = 3.5$

expected value → E; of 1 → 1; Roll → R

> These are synonyms:
> Expected Value = Mean = Average

2) After 100 rolls you can expect to win:

$$E_{100}(R) = 100\cdot E_1(R) = \$350$$

expected value → E; of 100 → 100; Rolls → R

Q: What will those be for a 60-face die?

Ex. 1b In the previous Example, how much should you expect to deviate from that average amount of \$350?

Sol'n: The answer here is much less intuitive, and therefore I only state the algorith.

1) Compute a quantity called **variance**:

$$\underbrace{V_1}_{\text{variance}}(\underbrace{R}_{\text{Roll of one}}) = (\underbrace{1-E_1(R))^2 \cdot p_1}_{\text{same}} + (\underbrace{2-E_1(R))^2 \cdot p_2}_{\text{same}} + \dots + (\underbrace{6-E_1(R))^2 \cdot p_6}_{\text{same}}$$

note that each term is ≥ 0 because of $(\dots)^2$

$$= \left[(1-3.5)^2 + (2-3.5)^2 + \dots + (6-3.5)^2\right] \cdot \frac{1}{6} \quad p_1 = p_2 = \dots = p_6$$

$$\approx 2.92$$

2) A measure of the spread of the results of one roll is given by the **standard deviation:**

Greek lower-case "sigma" $\rightarrow$ $\boxed{\sigma_1(R) = \sqrt{V_1(R)}}$

In our case: $\sigma_1(R) = \sqrt{2.92} \approx 1.71$.

To better understand its meaning, consider 100 independent rolls. Here, the situation is less intuitive than with the Expected Value in Ex. 1a.

Continuing with the algorithm:

3) variance of 100 rolls $\rightarrow$ $\underbrace{V_{100}(R) = 100 \cdot V_1(R)}_{\text{same}}$

$$\boxed{\sigma_{100}(R)} = \sqrt{V_{100}(R)} = \sqrt{100 \cdot V_1(R)} = \sqrt{100} \cdot \sqrt{V_1(R)} = \boxed{\sqrt{100}\, \sigma_1(R)}$$

The interpretation is that with "a high probability", after a 100 rolls you will win **between** $(E_{100}(R) - \sigma_{100}(R))$ **and** $(E_{100}(R) + \sigma_{100}(R))$, i.e. between $\$(350-17) \simeq \$333$ and $\$(350+17) \simeq \$367$.

Again, **the standard deviation characterizes the spread of values around the Expected Value (a.k.a. mean).**

Question: How "high" is the probability to find values of a random variable X between $(E(X) - \sigma(X),\ E(X) + \sigma(X))$?

Answer: We will consider a special case in Lec. 9.

Question: Why did we multiply the variance and not the standard deviation by the number of rolls (100)?

Answer: This is studied in courses on Probability & Statistics.

Conclusions for discrete variables

- **Expected value** (= mean = average) of a random variable R taking on values $r_1, r_2, \dots r_n$ with respective probabilities $p_1, p_2, \dots p_n$ is:

$$E(R) = r_1 \cdot p_1 + r_2 \cdot p_2 + \dots + r_n \cdot p_n$$

This means that if you observe R many times, then on average you will observe its value to be (near) $E(R)$.

- Variance of R:

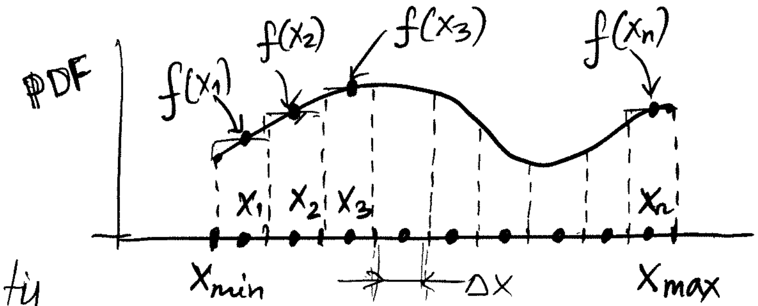
$$V(R) = (r_1 - E(R))^2 \cdot p_1 + (r_2 - E(R))^2 \cdot p_2 + \dots + (r_n - E(R))^2 \cdot p_n$$

- Standard deviation: $\sigma(R) = \sqrt{V(R)}$

The standard deviation characterizes how spread out individual observations of R are around $E(R)$.

② Continuous random variables

We know from Lec. 7 that if X is a continuous random variable, then probability p_i to find X in the i-th sub-interval between X_{min} and X_{max} (see the figure) is: $p_i \approx f(x_i)\Delta x$



Then the definitions of the Expected Value and Variance for a **continuous variable** X with the PDF $f(x)$ are:

$$\boxed{E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx}$$

← Expected Value = Mean = Average

MUST MEMORIZE !

$r_i \cdot p_i$ ← This is what it was in the discrete case.

Another notation: $\boxed{E(X) = \mu}$ ← Greek "mu"

$$\boxed{V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\, dx}$$

← Variance of X

The standard deviation is: $\boxed{\sigma(X) = \sqrt{V(X)}}$

Alternative formula for variance (proof on p. 643):

$$\boxed{V(X) = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx - \mu^2}$$

It is sometimes easier to use than the original definition.

MUST MEMORIZE!

Ex. 2 (= Ex. 11.3.1 in textbook)

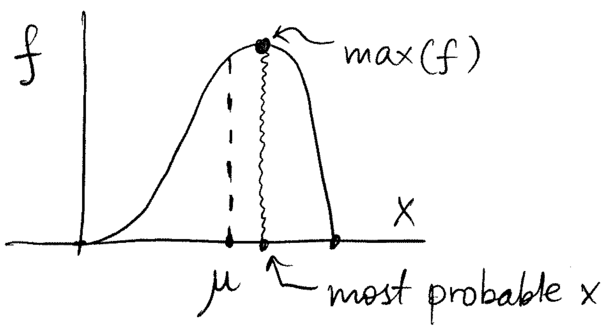
Find the expected value (= mean), variance, and standard deviation for the PDF $f(x) = \begin{cases} 12x^2 - 12x^3, & 0 \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$

Sol'n: 1) $\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx =$

$$= \int_{-\infty}^{0} x \cdot \cancelto{0}{f(x)}\,dx + \int_{0}^{1} x \cdot f(x)\,dx + \int_{1}^{\infty} x \cdot \cancelto{0}{f(x)}\,dx$$

$$= \int_0^1 x \cdot f(x)\,dx = \int_0^1 x \cdot (12x^2 - 12x^3)\,dx = \int_0^1 (12x^3 - 12x^4)\,dx$$

$$= \left(\frac{12}{4}x^4 - \frac{12}{5}x^5\right)\Big|_0^1 = \left(3 \cdot 1 - \frac{12}{5} \cdot 1\right) - 0 = \frac{3}{5} = 0.6$$



Interpertation:

If you observe the variable X with this PDF and record the values of your observation, than the average of your recorded values will be $\mu = 0.6$ (for many-many observations).

Note that this will not be your most frequently observed value. The latter corresponds to where $f(x)$ reaches its maximum and can be shown to be $\frac{2}{3} \approx 0.67$.

2) For the variance, we can choose between using either of the two formulas:

$$V(X) = \int_0^1 (x - \mu)^2 \cdot f(x)\,dx \qquad (1)$$

same bounds as before

or $\quad V(X) = \int_0^1 x^2 f(x)\,dx - \mu^2. \quad (2)$

If we use (1):

$$V(X) = \int_0^1 \left(x-\tfrac{3}{5}\right)^2 (12x^2 - 12x^3)\,dx =$$

$$= \int_0^1 \left(x^2 - 2\cdot\tfrac{3}{5}x + \left(\tfrac{3}{5}\right)^2\right)(12x^2 - 12x^3)\,dx = \ldots$$

This will require us to multiply two polynomials...
See Ex. 11.3.1 as done in the textbook.

But if we begin with the equivalent formula (2):

$$V(X) = \int_0^1 x^2 (12x^2 - 12x^3)\,dx - \left(\tfrac{3}{5}\right)^2 =$$

$$= \int_0^1 (12x^4 - 12x^5)\,dx - \left(\tfrac{3}{5}\right)^2 = \left(\tfrac{12}{5}x^5 - \tfrac{12}{6}x^6\right)\Big|_0^1 - \left(\tfrac{3}{5}\right)^2$$

$$= \left[\left(\tfrac{12}{5}\cdot 1 - \tfrac{12}{6}\cdot 1\right) - 0\right] - \tfrac{9}{25} = \tfrac{2}{5} - \tfrac{9}{25} = \tfrac{1}{25}.$$

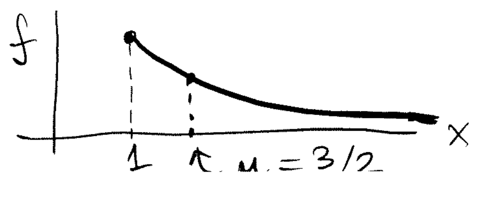
Eq. (2) is much easier to use! (Will not be so in the next Lecture, but here it is.)

3) $\sigma = \sqrt{V(X)} = \sqrt{\tfrac{1}{25}} = \tfrac{1}{5}.$

<u>Ex. 3</u> (= Ex. 11.3.4 in textbook)

Find the expected value, variance, and standard deviation of a random variable with PDF: $f(x) = \begin{cases} 3/x^4, & x \geq 1 \\ 0, & \text{otherwise}. \end{cases}$

<u>Sol'n:</u>



1) Expected value

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx =$$

$$= \int_1^\infty x \cdot \frac{3}{x^4}\,dx = \int_1^\infty \frac{3}{x^3}\,dx = \int_1^\infty 3x^{-3}\,dx$$

integrate only where PDF $\neq 0$

By Lec. 6, this will converge!

$$= 3\lim_{b\to\infty}\int_1^b x^{-3}\,dx = 3\lim_{b\to\infty} \left.\frac{x^{-3+1}}{-3+1}\right|_1^b =$$

$$= 3\lim_{b\to\infty}\left(\cancel{-\frac{1}{2b^2}}^{\to 0} - \frac{1}{-2\cdot 1^2}\right) = 3\left(0+\frac{1}{2}\right) = \frac{3}{2}.$$

<u>Note:</u> Again, $\mu = 3/2$ is <u>not</u> the most likely observed value of X. Rather, it will be the average of many-many observed values of X.

2) The previous example has taught us that it is probably easier to compute the variance using the "Alternative formula" (end of p. 8-4, or Eq. (2) on p. 8-6). So use it here:

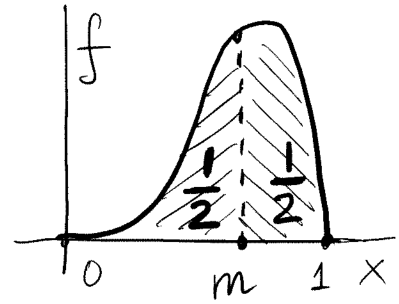
$$V(X) = \int_1^\infty x^2\cdot f(x)\,dx - \mu^2 = \int_1^\infty x^2\cdot\frac{3}{x^4}\,dx - \left(\frac{3}{2}\right)^2$$

$$= \int_1^\infty 3x^{-2}\,dx - \left(\frac{3}{2}\right)^2 = 3\left.\frac{x^{-2+1}}{-2+1}\right|_1^\infty - \frac{9}{4} =$$

$$= 3\left(\cancel{\frac{\infty^{-1}}{-1}}^{\to 0} - \frac{1^{-1}}{-1}\right) - \frac{9}{4} = 3(0+1) - \frac{9}{4} = \frac{3}{4}.$$

3) $\sigma = \sqrt{V(X)} = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2} \approx 0.87$

③ Median, and how it differs from mean



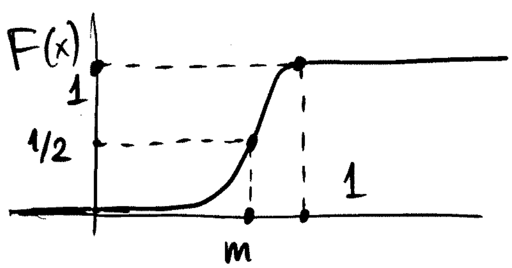
Recall that the total area under the PDF is 1. (In the figure we show the familiar PDF of Ex. 2.)
Then there must be point $x = m$ such that on the left and on the right of it the areas under the PDF curve are $\frac{1}{2}$.
In other words:

$$\boxed{P(X \le m) = \frac{1}{2} = P(X \ge m)}$$

This value, $x = m$, is called the **median**.

Ex. 4 Find the median of the PDF of Ex. 2

Sol'n: 1) Find the CDiF $= P(X \le x) = F(x)$.



We did this in Lec. 7, Ex. 3 (p. 7-13):

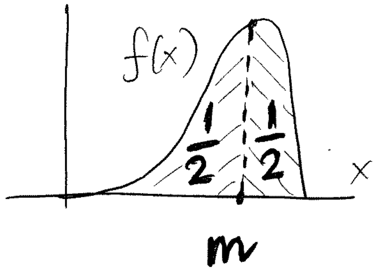
$$F(x) = \begin{cases} 0, & x < 1 \\ 4x^3 - 3x^4, & 0 \le x \le 1 \\ 1, & x > 1. \end{cases}$$

2) Solve the equation $P(X \le m) = 1/2$, or, equivalently:

$$F(m) = 1/2$$

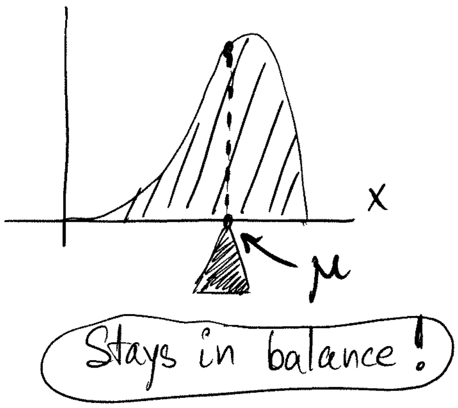
In this case: $4m^3 - 3m^4 = 1/2$.
This must be solved numerically: $m \approx 0.61$.
Note: It is close to the mean ($\mu = 0.60$, Ex. 2), but not the same.
(In general, mean and median may not be as close as in this example.)

Geometric interpretation of median:



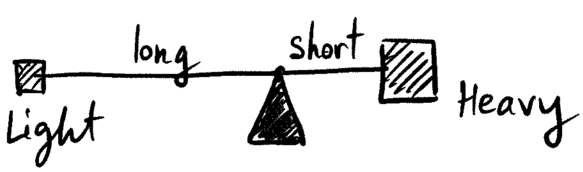
Divides the area under the PDF into two equal parts (i.e., into halves).

Geometric interpretation of mean:



- Imagine a plate cut out of a uniform metal sheet (or plywood) in the shape of the PDF.
- If you place a fulcrum (support) under $x = \mu$, then this plate will be in perfect balance.

- Note that this is not the same as saying that the **masses on the left and right of the fulcrum must be equal.** The balance also depends on where the left and right parts have their "effective centers". This is similar to the



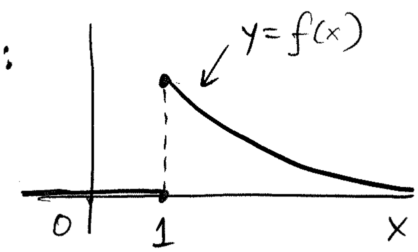
Lever Law:

A child on the longer arm of a swing can balance their parent sitting on the shorter arm.

Ex. 5 (≃ Ex. 11.3.5 in textbook)

Find the median of the PDF of Ex. 3:

$$f(x) = \begin{cases} 3/x^4, & x \geq 1 \\ 0, & \text{otherwise.} \end{cases}$$



Sol'n: Follow the steps of Ex. 4.

1) Find the CDiF $= P(X \le x) = F(x)$.

$\underline{x<1}$: $F(x) = \int_{-\infty}^{x} \cancel{f(t)}^{\,0}\, dt = 0$

$\underline{x \ge 1}$: $F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{1} \cancel{f(t)}^{\,0}\,dt + \int_{1}^{x} f(t)\,dt$

$= \int_{1}^{x} 3 \cdot t^{-4}\,dt = 3 \cdot \frac{t^{-4+1}}{-4+1}\Big|_1^x = 3\left(\frac{x^{-3}}{-3} - \frac{1^{-3}}{-3}\right)$

$= 1 - \frac{1}{x^3}$.

2) Solve the equation $P(X \le m) = \frac{1}{2}$, or $F(m) = \frac{1}{2}$:

$1 - \frac{1}{m^3} = \frac{1}{2} \Rightarrow 1 - \frac{1}{2} = \frac{1}{m^3} \Rightarrow$

$\frac{1}{2} = \frac{1}{m^3} \Rightarrow m^3 = 2 \Rightarrow \boxed{m = \sqrt[3]{2} \approx 1{,}26}$